A Data Quality Screening Service for Remote Sensing Data

Christopher Lynnes, NASA/GSFC

Edward Olsen, NASA/JPL

Peter Fox, RPI

Bruce Vollmer, NASA/GSFC

Robert Wolfe, NASA/GSFC

***Contributions from**: N. Most, Y.-I. Won, T. Hearty, R. Strub, S. Ahmad, S. Zednik, M. Hegde and A. Rezaiyan-Nojani*

***Funded by**: NASA's Advancing Collaborative Connections for Earth System Science (ACCESS) Program*

Outline

- Why a Data Quality Screening Service?
- Making Quality Information Easier to Use via the Data Quality Screening Service (DQSS)
- Future Directions for DQSS

Why a Data Quality Screening Service?

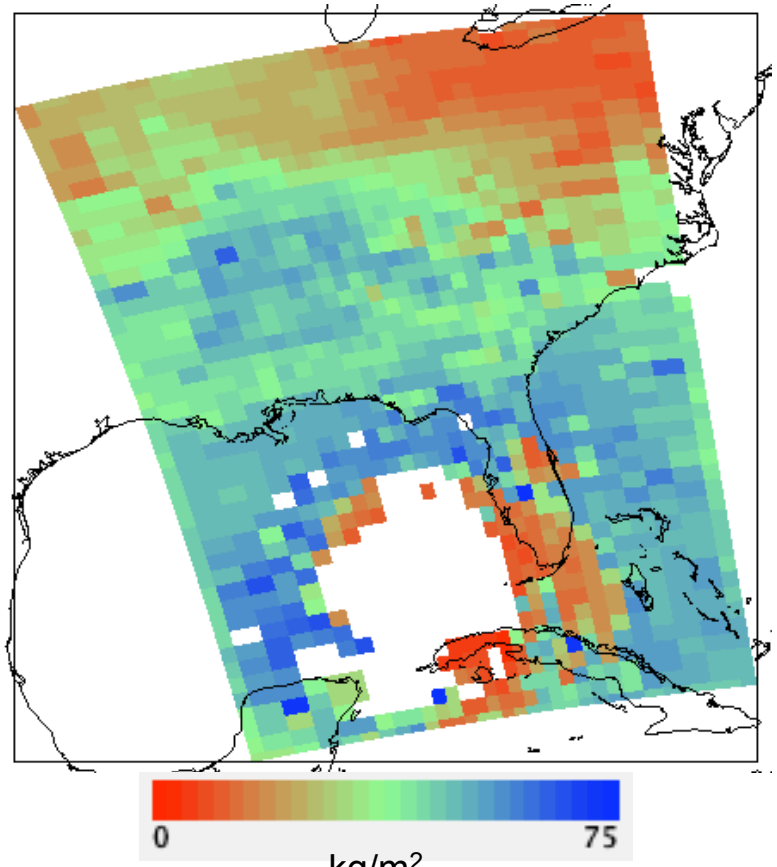
The quality of AIRS data varies considerably

AIRS Parameter	Best (%)	Good (%)	Do Not Use (%)
Total Precipitable Water	38	38	24
Carbon Monoxide	64	7	29
Surface Temperature	5	44	51

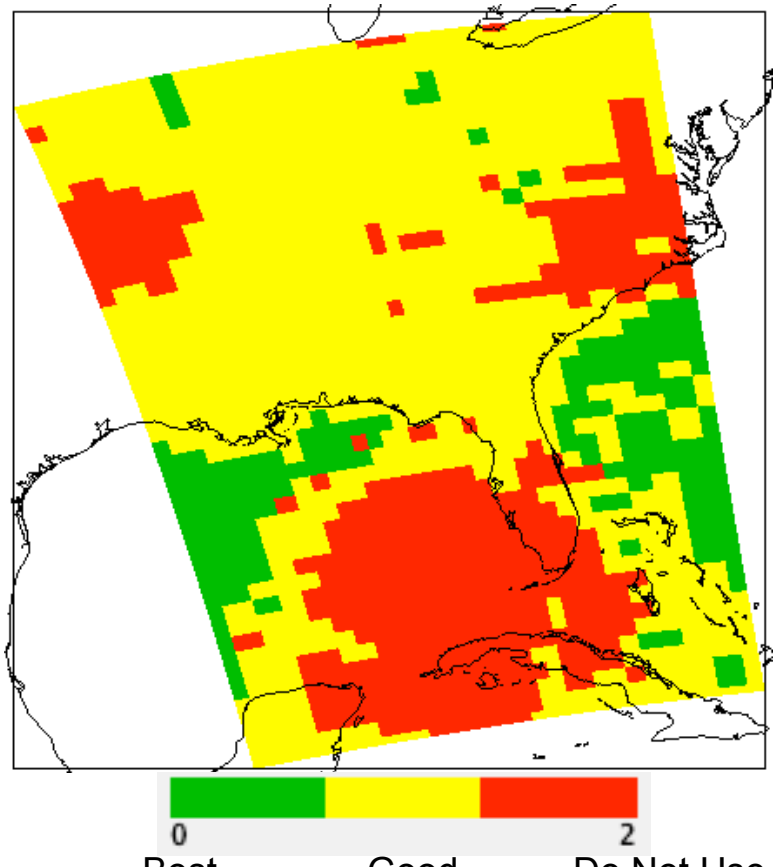
Version 5 Level 2 Standard Retrieval Statistics

Quality schemes can be relatively simple…

Total Column Precipitable Water

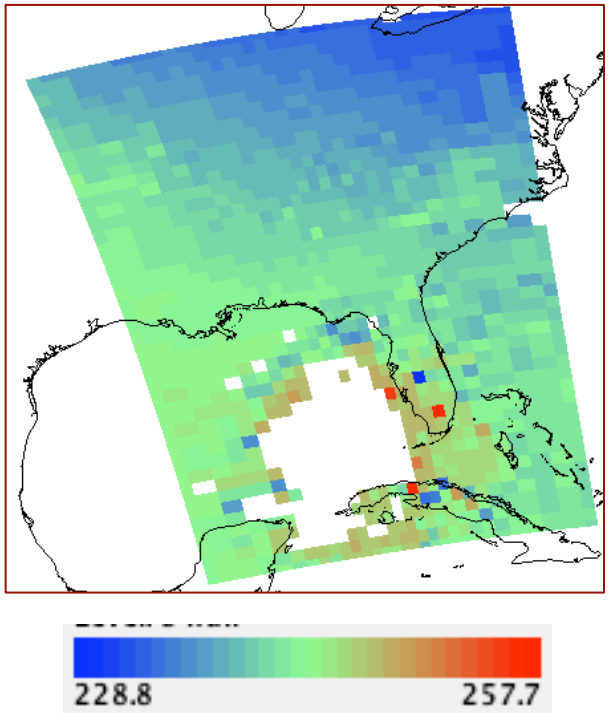


Qual_H2O

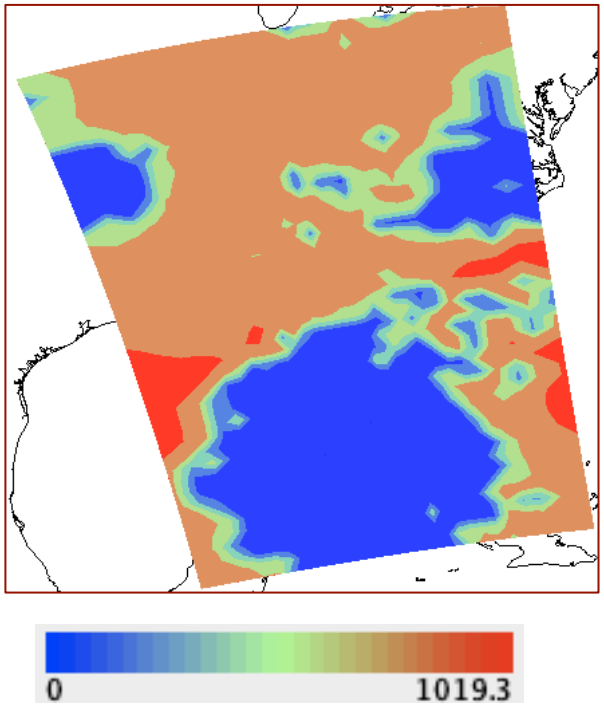


…or they can be more complicated

Hurricane Ike, viewed by the Atmospheric Infrared Sounder (AIRS)

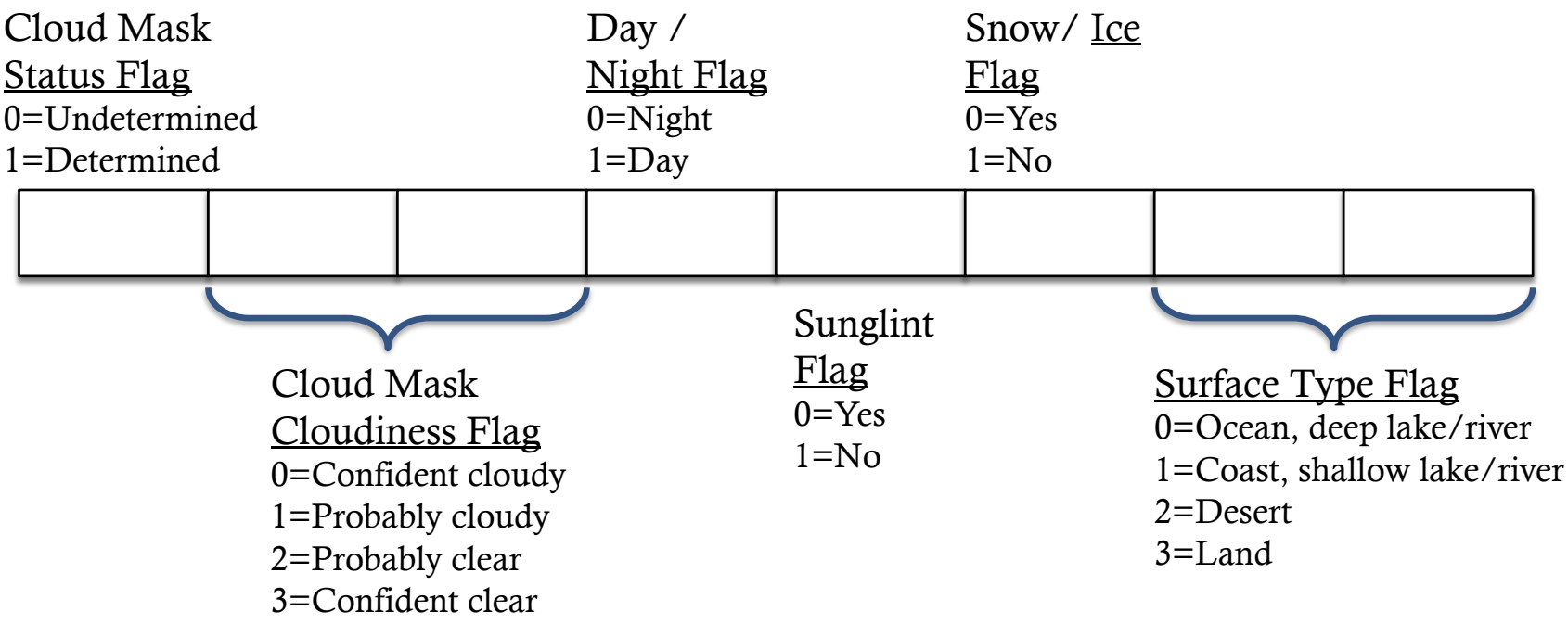


Air Temperature
at 300 mbar



PBest : Maximum pressure for which quality value is “Best” in temperature profiles

Quality flags are also sometimes packed together into bytes



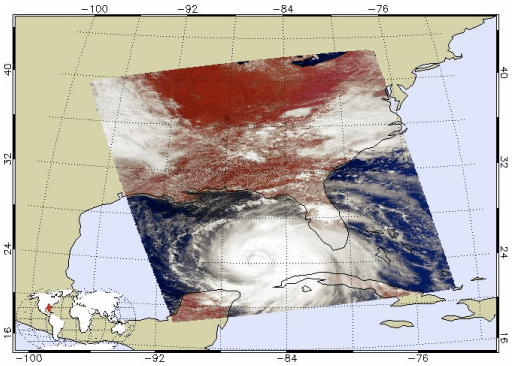
Bitfield arrangement for the Cloud_Mask_SDS variable in atmospheric products from Moderate Resolution Imaging Spectroradiometer (MODIS)

Current user scenarios…

- Nominal scenario
  - Search for and download data
  - Locate documentation on handling quality
  - Read & understand documentation on quality
  - Write custom routine to filter out bad pixels

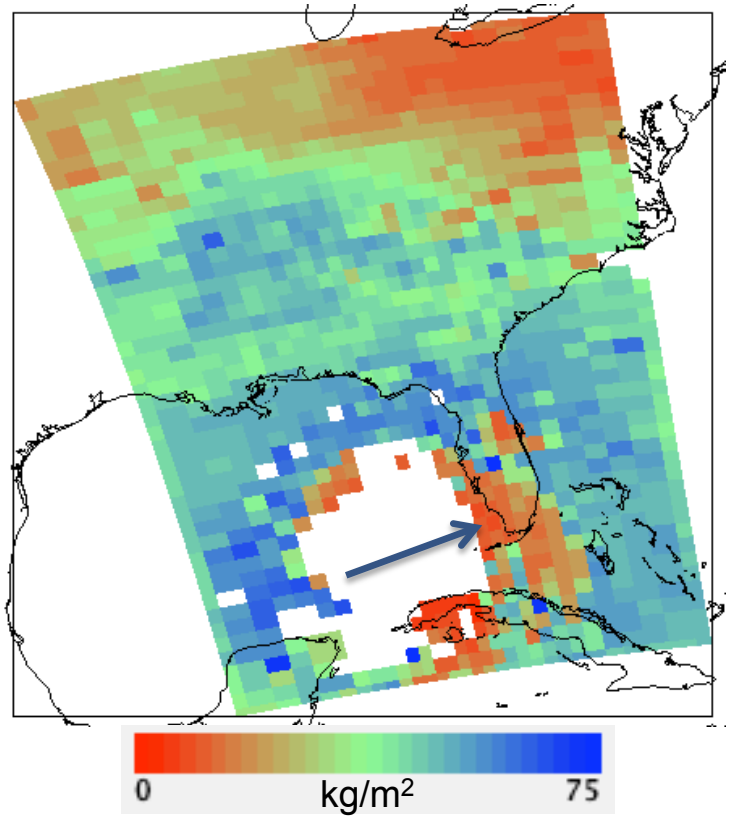
  Repeat for each user
- Equally likely scenario (*especially in user communities not familiar with satellite data*)
  - Search for and download data
  - Assume that quality has a negligible effect

The effect of bad quality data is often *not* negligible

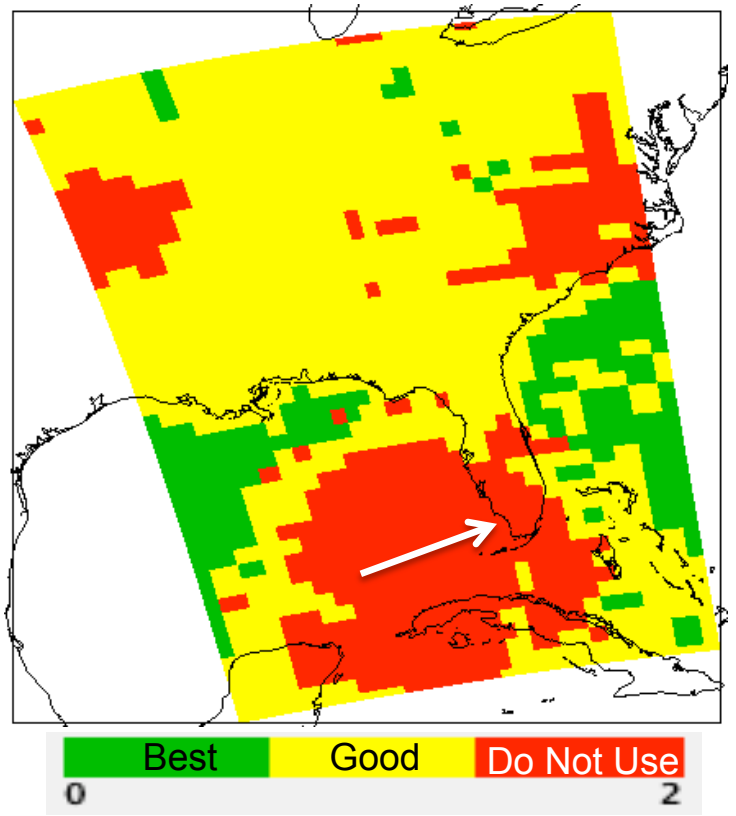


Hurricane Ike, 9/10/2008

Total Column Precipitable Water



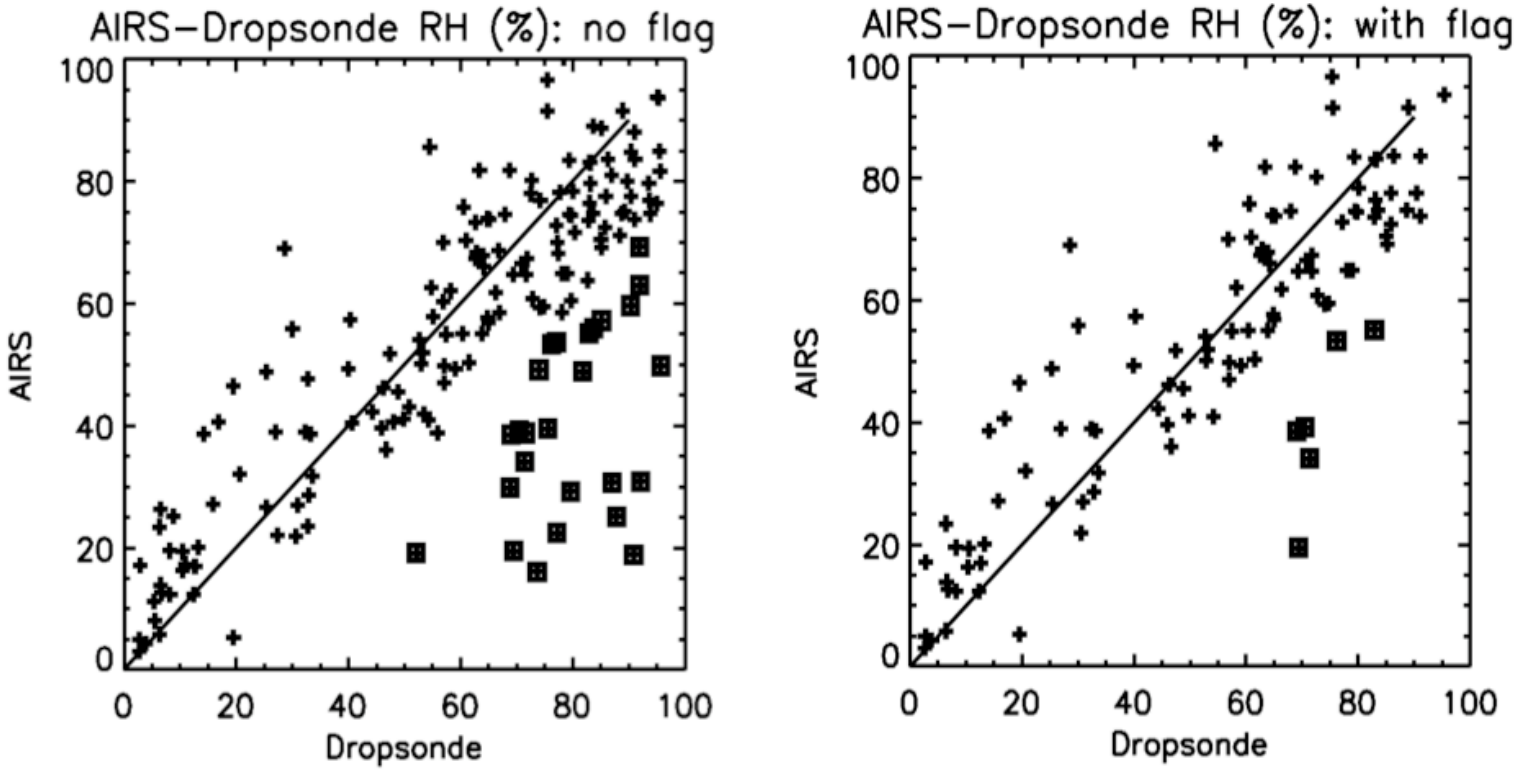
Quality



Neglecting quality may introduce bias (a more subtle effect)

AIRS Relative Humidity Comparison against Dropsonde
with and without Applying PBest Quality Flag Filtering

Boxed data points indicate AIRS RH data with dry bias > 20%



From a study by Sun Wong (JPL) on specific humidity in the Atlantic Main Development Region for Tropical Storms

Making Quality Information Easier to Use via the Data Quality Screening Service (DQSS)

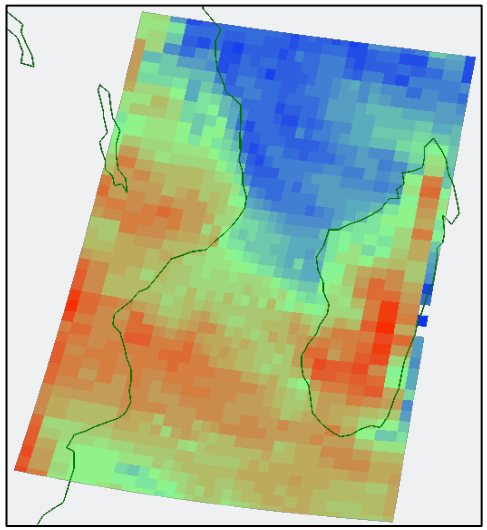
.

The DQSS filters out bad pixels for the user

- Default user scenario
  - Search for data
  - Select science team recommendation for quality screening (filtering)
  - Download screened data
- More advanced scenario
  - Search for data
  - Select custom quality screening parameters
  - Download screened data

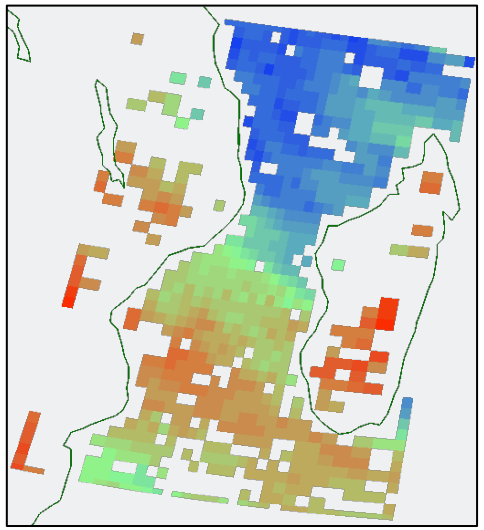
DQSS replaces bad-quality pixels with fill values

Original data array (Total column precipitable water)



Mask based on user criteria (Quality level < 2)

Good quality data pixels retained



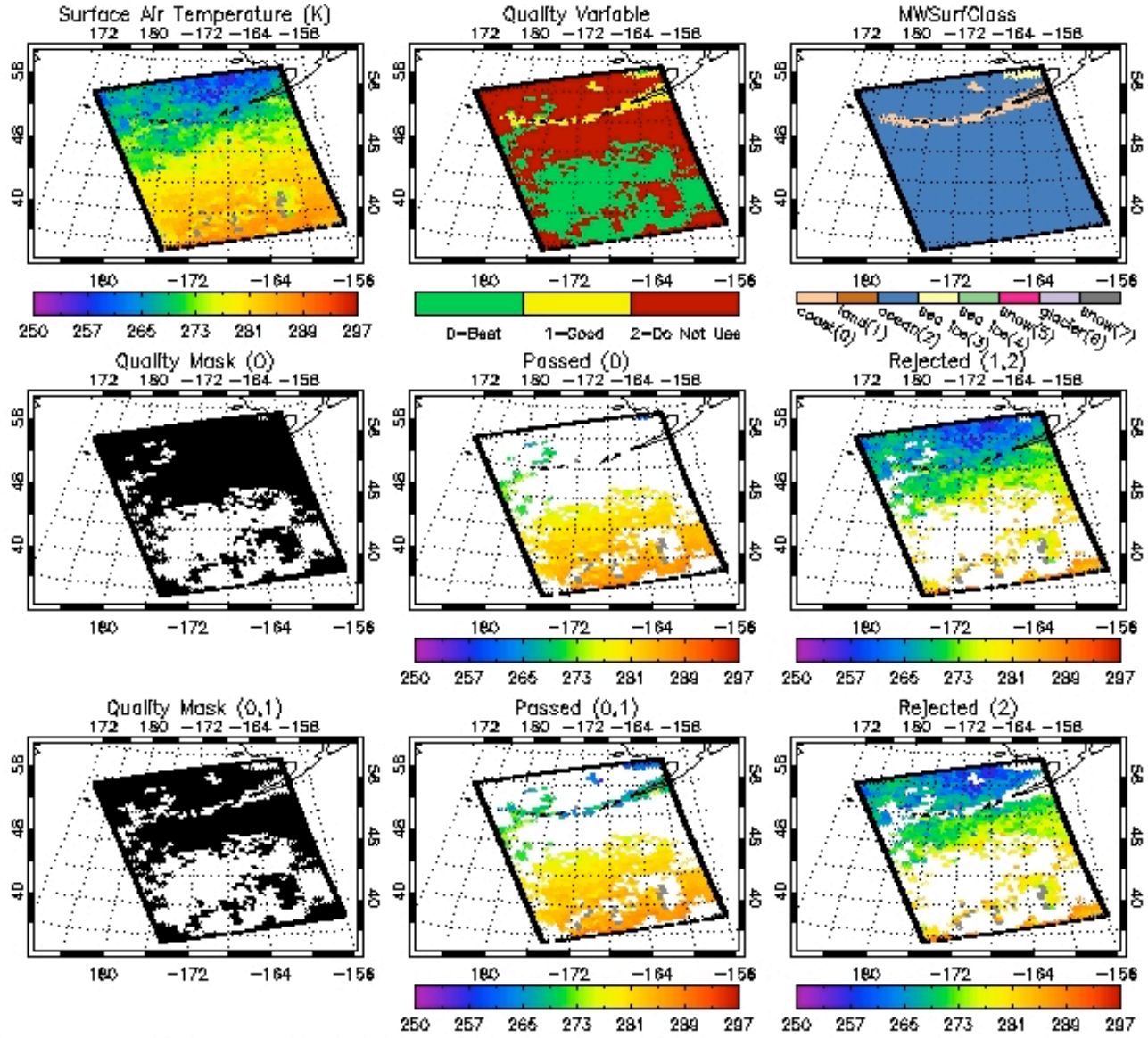
Output file has the same format and structure as the input file (except for extra mask and original_data fields)

Visualizations help users see the effect of different quality filters

Data within product file

Best quality only

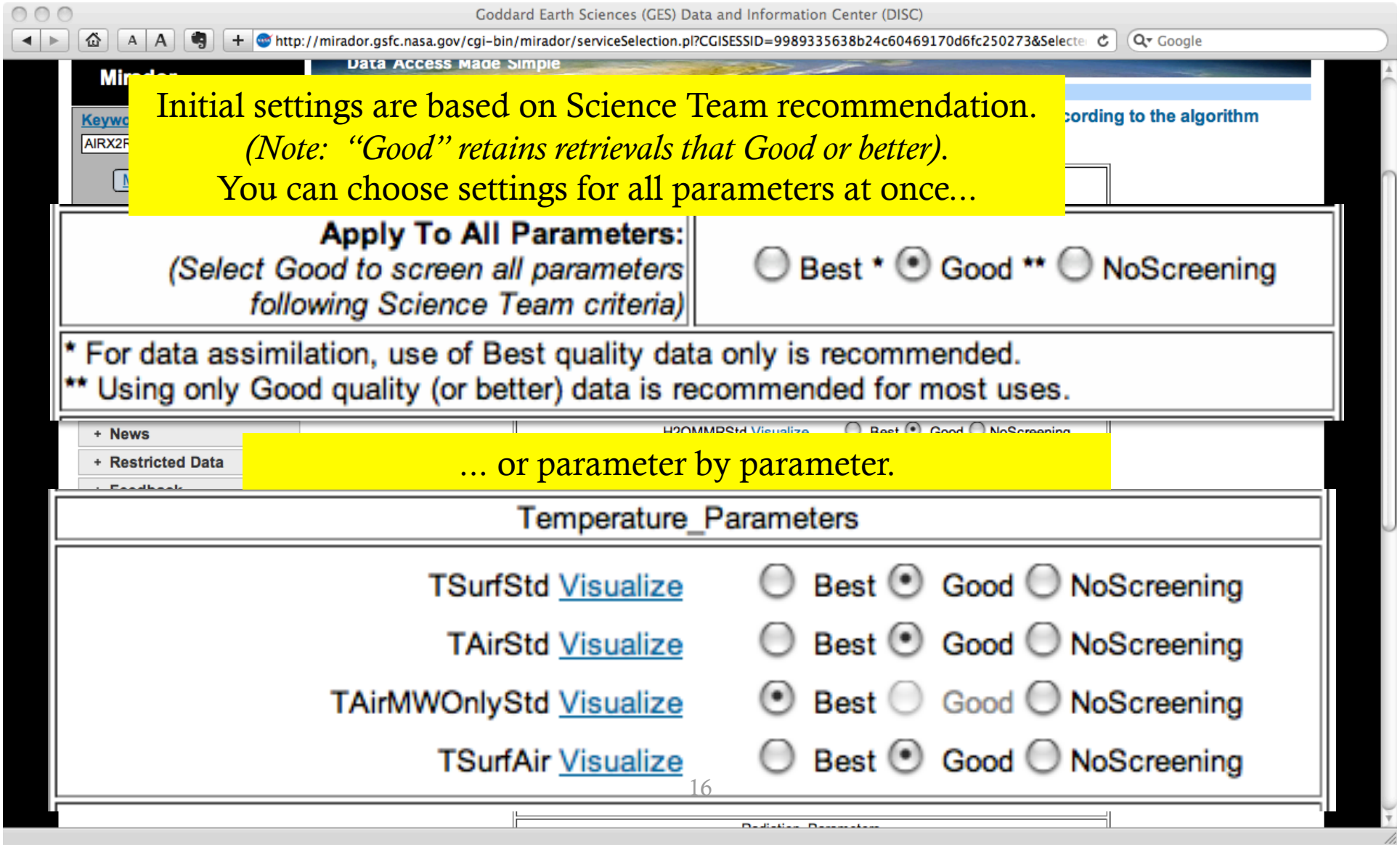
Best + Good quality



DQSS can encode the science team recommendations on quality screening

- *AIRS Level 2 Standard Product*
  - Use Best-only for data assimilation uses
  - Use Best+Good for climatic studies
- *MODIS Aerosols*
  - Use only VeryGood over land
  - Use Marginal+Good+VeryGood over ocean

Or, users can select their own criteria…



Status and Future

DQSS Status

- DQSS is operational for AIRS L2 Standard Products
  - DQSS is offered through the Mirador data search interface at the GES DISC
- Usage Metrics will be collected:
  - Basic usage
  - What criteria are used
    - Screening invocation is a simple URL GET
  - What the " yield" was for the screening (future)

Future: Extend DQSS to Other Datasets

- Moderate Resolution Imaging Spectroradiometer (MODIS) L2 Atmospheres
- Microwave Limb Sounder (MLS) L2
- Ozone Monitoring Instrument (OMI) L2
- High-Resolution Dynamics Limb Sounder (HIRDLS)

Combine DQSS with Other Services

- Subsetting
- Reformatting to NetCDF
- OPeNDAP
- Suggestions?

DQSS Target Uses and Users

	Routine	Visual-ization	Quick Recon.	Machine-level	Metrics
Interdisciplinary	X	X			
Educational	X	X			
Expert/Power	?		X	X	
Applications				X	
Algorithm Developers					X

DQSS Recap

- Screening satellite data can be difficult and time consuming for users
- The Data Quality Screening System will provide an easy-to-use service
- The result should be:
  - More attention to quality on users' part
  - More accurate handling of quality information…
  - …With less user effort